COVER SHEET FOR PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION

PROGRAM ANNOUNCEMENT/SOLICITATION NO./CLOSING DATE/if not in response to a program announcement/solicitation enter NSF 11-1

NSF 11-545 **09/06/11**

FOR CONSIDERATION BY NSF ORGANIZATION UNIT(S) (Indicate the most specific unit known, i.e. program, division, etc.)

MCB - Genetic Mechanisms

FOR NSF USE ONLY

NSF PROPOSAL NUMBER

1158119

DATE RECEIVED	NUMBER OF COPIES	DIVISION ASSIGNED	FUND CODE	DUNS# (Data Universal Numbering System)	FILE LOCATION
09/06/2011	1	08070000 MCB	1112	605799469	02/28/2012 1:10pm S

EMPLOYER IDENTIFICATION NUMBER (EIN) OR TAXPAYER IDENTIFICATION NUMBER (TIN)

916001537

SHOW PREVIOUS AWARD NO. IF THIS IS
☐ A RENEWAL
☐ AN ACCOMPLISHMENT-BASED RENEWAL

IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY? YES ☐ NO ☒ IF YES, LIST ACRONYM(S)

NAME OF ORGANIZATION TO WHICH AWARD SHOULD BE MADE

University of Washington

AWARDEE ORGANIZATION CODE (IF KNOWN)

0037986000

ADDRESS OF AWARDEE ORGANIZATION, INCLUDING 9 DIGIT ZIP CODE

SEATTLE, WA 981959472
US

NAME OF PRIMARY PLACE OF PERF

University of Washington

ADDRESS OF PRIMARY PLACE OF PERF, INCLUDING 9 DIGIT ZIP CODE

University of Washington
1140 NE Boat Street
Seattle ,WA ,981056692 ,US.

IS AWARDEE ORGANIZATION (Check All That Apply) (See GPG II.C For Definitions)
☐ SMALL BUSINESS ☐ MINORITY BUSINESS ☐ FOR-PROFIT ORGANIZATION ☐ WOMAN-OWNED BUSINESS

☐ IF THIS IS A PRELIMINARY PROPOSAL THEN CHECK HERE

TITLE OF PROPOSED PROJECT **DNA methylation as a mechanism to increase adaptive potential in invertebrates**

REQUESTED AMOUNT	PROPOSED DURATION (1-60 MONTHS)	REQUESTED STARTING DATE	SHOW RELATED PRELIMINARY PROPOSAL NO. IF APPLICABLE
$ 243,090	24 months	03/01/12	

CHECK APPROPRIATE BOX(ES) IF THIS PROPOSAL INCLUDES ANY OF THE ITEMS LISTED BELOW

☐ BEGINNING INVESTIGATOR (GPG I.G.2)
☐ DISCLOSURE OF LOBBYING ACTIVITIES (GPG II.C.1.e)
☐ PROPRIETARY & PRIVILEGED INFORMATION (GPG I.D, II.C.1.d)
☐ HISTORIC PLACES (GPG II.C.2.j)
☐ EAGER* (GPG II.D.2) ☐ RAPID** (GPG II.D.1)
☐ VERTEBRATE ANIMALS (GPG II.D.6) IACUC App. Date _____
PHS Animal Welfare Assurance Number _____
☐ HUMAN SUBJECTS (GPG II.D.7) Human Subjects Assurance Number _____
Exemption Subsection _____ or IRB App. Date _____
☐ INTERNATIONAL COOPERATIVE ACTIVITIES: COUNTRY/COUNTRIES INVOLVED (GPG II.C.2.j)
_____
☐ HIGH RESOLUTION GRAPHICS/OTHER GRAPHICS WHERE EXACT COLOR REPRESENTATION IS REQUIRED FOR PROPER INTERPRETATION (GPG I.G.1)

PI/PD DEPARTMENT
Aquatic & Fishery Sciences

PI/PD FAX NUMBER

PI/PD POSTAL ADDRESS
4333 Brooklyn Ave NE

SEATTLE, WA 981959472
United States

NAMES (TYPED)	High Degree	Yr of Degree	Telephone Number	Electronic Mail Address
PI/PD NAME **Steven Roberts**	**PhD**	**2002**	**206-543-4043**	**sr320@u.washington.edu**
CO-PI/PD				
CO-PI/PD				
CO-PI/PD				
CO-PI/PD				

Electronic Signature

DNA methylation as a mechanism to increase adaptive potential in invertebrates
PI: Steven Roberts, University of Washington

Intellectual Merit

The study of DNA methylation is providing remarkable insight into gene regulation and the complex mechanisms associated with phenotypic variation and adaptation to environmental change. DNA methylation is present in species from prokaryotes to humans, however there is dramatic diversity in characteristics - from species where this phenomenon is absent, to cases where the genome is globally methylated. Although DNA methylation is considered an evolutionarily ancient epigenetic mechanism, the role and evolutionary significance of this process is not fully understood. The overall goal of the research is to use the Pacific oyster, *Crassostrea gigas*, as a model system to evaluate the potentially transformative theory that ***the absence of germline methylation in genes involved in adapting to changing conditions has evolved to increase adaptive potential in organisms exposed to heterogeneous environments.*** The Pacific oyster is an excellent model to use for the proposed project because it is one of only a few invertebrates where DNA methylation has been described, there are considerable genomic resources available, and the Pacific oyster is a key bioindicator species that can provide us with a better understanding of how species respond to fluctuating environmental conditions.

The specific research objectives are to 1) experimentally characterize the DNA methylation landscape in *Crassostrea gigas* and 2) evaluate the relationship between methylation, gene expression, alternative splicing, and sequence mutations. In order to carry out these objectives ***a combination of high-throughput sequencing and DNA tiling array analysis will be coupled with methylation enrichment to determine genome wide methylation distribution and assess how specific DNA methylation patterns influence transcriptional activity***. Multi-generational samples will be examined as well as individual oysters prior to and following an environmental perturbation (*i.e.* temperature change). The experimental design will allow for a comprehensive characterization of DNA methylation including evaluating the theory that genes lacking germline methylation will possess a greater incidence of alternative transcripts, sequence variation, and transient methylation.

Broader Impact

The research results from this project will provide significant genomic resources that will contribute to the understanding of how genes work and genomes change over time. Furthermore, a characterization of the functional role of DNA methylation in this lineage will provide important insight into the evolution of DNA methylation. The basis of the proposed theory suggests DNA methylation will play a critical role in the ability of species to respond to alterations associated with large-scale environmental change by increasing the probability of successful adaptations. ***Thus, this research could considerably impact our predictions on ecosystem level responses to climate change and ocean acidification***. This project will train and educate several students including two graduate students and undergraduate students. The research activities will also be integrated into the formal classroom setting. Research activities will be made available to the public and other scientists through open access electronic lab notebooks and as part of hands-on education and outreach activities. The public will gain an increased appreciation of molecular processes and their relationship to environmental change. This will be facilitated by the ability of the public to relate to the oyster, a species prominent in coastal ecosystems, aquariums, and markets.

Keywords: epigenetics, DNA methylation, gene expression, phenotypic plasticity, adaptation to environmental change, invertebrates

TABLE OF CONTENTS

For font size and page formatting specifications, see GPG section II.B.2.

*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

BACKGROUND AND OVERVIEW OF CURRENT KNOWLEDGE

DNA methylation is an epigenetic mark that can alter transcriptional activity. DNA methylation is present in species from prokaryotes to humans, however there is dramatic diversity in characteristics - from species where this phenomenon is absent, to cases where the genome is globally methylated. Likewise, the associated functional roles and mechanisms of methylation are predicted to be just as diverse. DNA methylation appears to be an ancient mechanism that has been co-opted to perform a variety of functions in different taxonomic lineages. However, despite its long history, the role and evolutionary significance of this epigenetic process is not fully understood. Our working theory postulates that DNA methylation increases adaptive potential in species commonly exposed to heterogeneous environmental conditions. Specifically, ***the absence germline methylation in genes involved in adapting to changing conditions facilitates the production of alternative transcripts and increases sequence variability***. Here we propose to characterize the relationship between DNA methylation and transcriptional activity in an intertidal mollusc, the Pacific oyster (*Crassostrea gigas*). ***The Pacific oyster is an excellent system to use for the proposed project because this species*** 1) ***is one of only a handful of invertebrates where DNA methylation has been described***, 2) ***has a significant amount of genomic resources available***, and 3) ***is a key bioindicator that can provide us with a better understanding of adaptation to environmental change***. In the remainder of this section we will provide background information on DNA methylation, with particular emphasis on invertebrates and the influence of environmental conditions on DNA methylation. Results from our ongoing research efforts characterizing DNA methylation in *C. gigas* will be presented, followed by an explanation of our working theory describing how DNA methylation influences transcriptional activity and associated phenotypic plasticity.

DNA Methylation Landscapes

DNA methylation is presumed to be evolutionarily ancient, and while the mark itself is prevalent among taxa there is an incredibly diverse landscape of methylation including differences in which residues are methylated, the density of the methylation, and the distribution of methylation across the genome. In animals, the primary targets of methylation are cytosines located 5' to guanines (CpG) in the DNA. Plant genomes can also be methylated at CpG sites as well as CpNpG and CpNpN sites, where N = A, C or T. Interestingly, recent studies have shown that human embryonic stem cells also contain a substantial amount of non-CpG methylation in the context of CpNpG and CpNpN comprising approximately 25% of all methylated cytosines (Lister et al. 2009).

The density of DNA methylation also varies significantly across taxa. In vertebrates, approximately 70-80% of cytosines in CpG dinucleotides are methylated (Bird and Taggart 1980), a pattern referred to as global methylation. In contrast, invertebrates display a wide range of DNA methylation. ***In fact, two of the most common model organisms (Drosophilia melanogaster and Caenorhabditis elegans) essentially lack DNA methylation*** (Gowher et al. 2000, Simpson et al. 1986). Likewise, genes coding for DNA methyltransferases have been partially (*D. melanogaster*) or completely lost (*C. elegans*) in these species. Other invertebrates have an intermediate level of methylation (30 – 60%) including sea urchins (*Strongylocentrotus purpuratus*) (Bird et al. 1979), *Ciona intestinalis* (Simmen and Bird 2000, Suzuki et al. 2007) and honey bees (*Apis mellifera*) (Lyko et al. 2010). Among plants, not all species studied have methylated genomes, and related species can exhibit varying degrees. For example, a global methylation pattern is observed in maize (*Zea mays*) (Palmer et al. 2003), whereas an intermediate level, similar to that seen in invertebrates, has been reported for *Arabidopsis thaliana* (Zhang et al. 2006).

The location of DNA methylation across the genome is diverse and related to density. In vertebrates, the limited amount of the genome that is not methylated is often found in CpG rich gene promoter regions called CpG islands. Gene bodies are typically methylated in vertebrates though the degree of methylation decreases in 5' and 3' regions. In invertebrates, tracts of methylated CpGs are interspersed with unmethylated regions, referred to as a mosiac pattern (Suzuki et al. 2007). Another example of spatial heterogeneity can be found in honey bees and *C. intestinalis* (Lyko et al. 2010, Suzuki et al. 2007) where methylation is predominantly found in exons, compared to the blood fluke (*Schistosoma mansoni*) where highly repetitive intronic regions are methylated (Geyer et al. 2011). In plants, methylation occurs predominantly on repetitive DNA elements and transposons (Zhang et al. 2006), though gene bodies are substantially methylated in some species (Zhang et al. 2006; Zilberman et al. 2006).

DNA Methylation as a Dynamic Process

DNA methylation is typically considered a stable epigenetic mark and established methylation patterns are faithfully maintained during mitosis by the enzyme DNA methyltransferase 1 (DNMT1), which preferentially methylates hemi-methylated DNA. Indeed, it is critical that certain DNA methylation states be maintained as in the case of imprinted genes and X chromosome inactivation in mammals. However, DNA methylation states can also vary in accordance with DNA replication, development, and in response to environmental conditions. Human DNMT1 has an error rate of 5% per CpG per cell division (Riggs et al. 1998), which can result in variations in methylation within the same cell type. Additionally, human CpG islands exhibit an increase in methylation during aging (*e.g.* Ahuja and Issa 2000), and in abnormal cells such as cancer (Cheung et al. 2009). While vertebrate methylation is considered important for development and tissue specific phenotypes, in *C. intestinalis* methylation patterns of sperm, embryos, and muscle tissue were indistinguishable (Suzuki et al. 2007).

Behavior can affect DNA methylation and has been demonstrated to be a mechanism involved in environmental programming. For example, infant rats deprived of grooming/licking by their mothers show an alteration in methylation in the promoter region of a glucocorticoid receptor. This change is correlated with increased anxiety behavior in the rats once they reach maturity (Meaney and Szyf 2005). Similarly, nutrition affects DNA methylation and subsequent phenotypes in bees. Bees fed a diet of royal jelly in early life shift development to a queen phenotype and DNA methylation is the primary mechanism by which royal jelly acts on the genome (Kucharski et al. 2008).

In plants, it has been proposed that variation in genome methylation may function as an adaptive mechanism to cope with environmental stress (reviewed by Chinnusamy and Zhu 2009). Dynamic methylation and demethylation in response to various stressors has been shown in a number of species. For instance, maize exposed to cold stress exhibits demethylation at nucleosome cores (Steward et al. 2002). In the common snapdragon (*Antirrhinum majus*), abiotic stresses such as aluminum, salt, and low temperature induce demethylation in the coding sequence of a glycerophospodiesterase, which is also correlated with induced gene expression (Choi and Sano 2007).

A Conserved Role of DNA Methylation?

If one considers the significant diversity of methylation across taxa it seems plausible that these marks could have different functions and potentially different mechanisms of action across evolutionary time. In fact, there have been several hypotheses regarding the evolution of DNA methylation based on the degree of methylation. An early hypothesis by Bestor

(1990) suggested that eukaryotic methylation evolved from bacterial methylation which functions primarily to recognize non-self DNA as a defense mechanism against bacteriophage. Based on evidence of homology between bacterial and eukaryotic DNA methyltransferases, Bestor hypothesized that the role of genome defense was expanded to regulate the function of the larger genomes in higher eukaryotes. A second hypothesis was proposed to explain the loss of methylation in certain species such as *Drosophilia.* This hypothesis suggested that organisms with short lifespans and low cell turnover make less use of the methylation system as a mechanism to support cell memory and subsequently lose methylation over evolutionary time (Jablonka and Lamb 1995). As we begin to learn more about DNA methylation thanks to advances in technology it is evident that these early hypotheses fail to address recent findings (see preliminary results below).

In the mammalian DNA methylation system, the conventional understanding is that DNA methylation is associated with transcriptional repression. The primary mechanism being that methylation inhibits binding of necessary promoter elements in the 5' UTR. This model is not suited for taxa with a mosaic methylation pattern where methylation is primarily found within coding sequences and not in promoter regions. ***Therefore, the discovery that the mosaic DNA methylation pattern was directly associated with gene function was a paradigm shift from the traditional mammalian model and likely holds the key with respect to a functional role and epigenetic mechanisms of intragenic DNA methylation***.

Some of the first evidence supporting a regulatory role of intragenic DNA methylation in invertebrates comes from computational analyses of the methylation status of the honey bee (*A. mellifera*) genome using *in silico* analysis to predict the methylation status of genes (Elango et al. 2009, Foret et al. 2009). This analysis is based on the known hyper-mutability of methylated cytosines, which readily deaminate to thymine residues (Coulondre et al. 1978). This CpG mutation is not easily corrected by DNA repair machinery and, as a result, consistently methylated regions of DNA are depleted of CpG dinucleotides over evolutionary time (Schorderet and Gartler 1992). Consequently, regions of DNA with a low CpG observed versus expected ratio (denoted as CpG O/E) are predicted to be methylated at the germline, whereas regions with a high CpG O/E (approaching 1.0) are predicted to be sparsely methylated. See **Figure 1** for an illustration of this concept. In honey bees, ubiquitously expressed, critical genes were predicted to be methylated at the germline, whereas caste-specific genes were predicted to lack methylation. From their study, Elango et al. (2009), hypothesized that genes predicted to be methylated (caste-specific) might have greater epigenetic flexibility, which allows for higher regulatory control of these inducible classes of genes via transient methylation.

More recently, direct measurements have been used to evaluate methylation in the honey bee genome. Using bisulfite treatment coupled with high-throughput sequencing, Lyko et al. (2010) found that methlyated cytosines occur primarily in exons and that methylated genes had a higher degree of conservation across species than unmethylated genes. ***These results confirmed an inverse relationship between germline methylation and CpG O/E.*** Other trends that arose from this analysis were that 1) methylated cytosine clusters were associated with alternatively spliced exons and 2) genes containing introns were more likely to be methylated than those lacking introns. The authors also highlighted an example where an increased level of methylation in an alternatively spliced exon in the worker bee brain was associated with an increased expression of the variant lacking the exon. Lyko et al. (2010) concluded methylation may not be functioning as an 'on/off' switch but instead allowing for '*fine tuning*' of transcriptional control of these conserved genes, including differentially spliced variants.

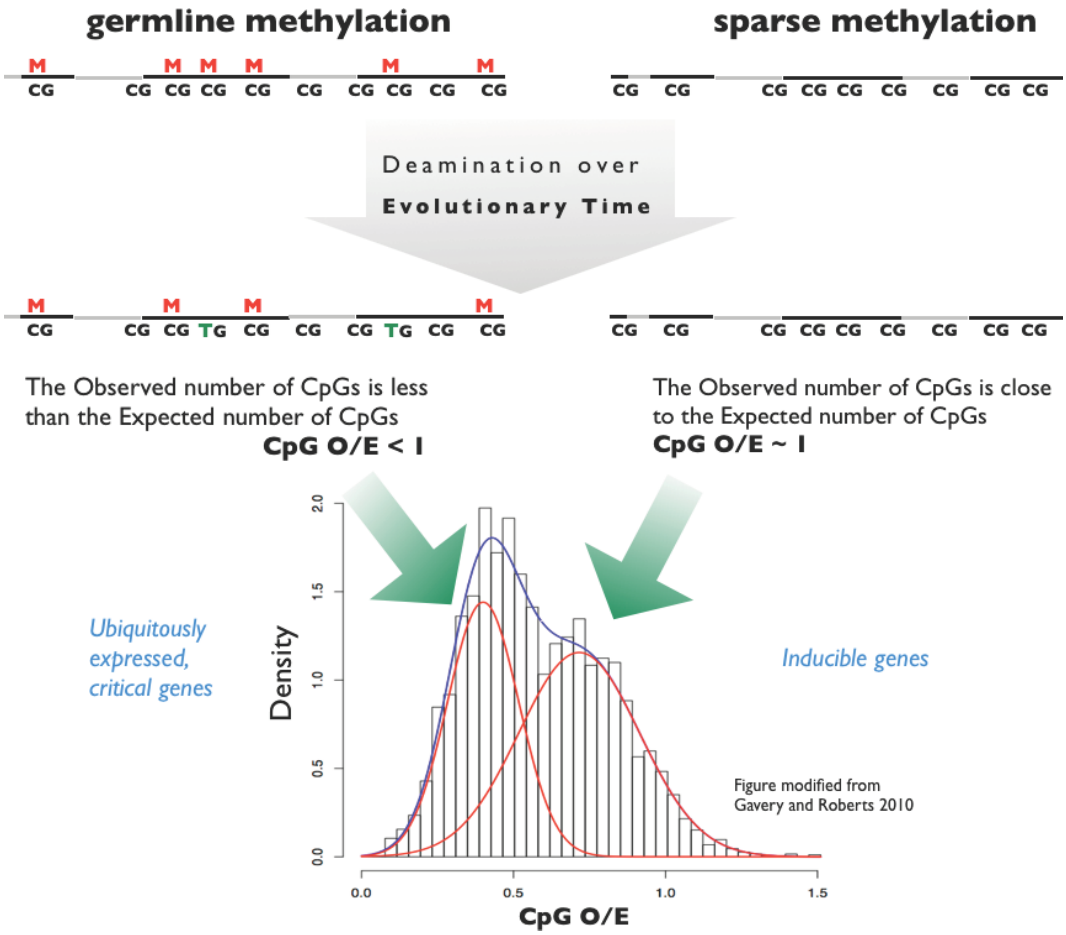

Figure 1. Illustration of biological underpinnings of CpG Observed versus Expected ratio (CpG O/E) used to identify portions of the genome that possess germline methylation. Over evolutionary time, methylated cytosines deaminate to thymine residues that are not easily corrected by proofreading machinery. Methylated regions will gradually lose CpG residues over time and thus the CpG O/E is negatively correlated with levels of DNA methylation. The histogram at the bottom of the figure is modified from Gavery and Roberts 2010 and represents the probability density function of CpG O/E for 12,210 *C. gigas* genes. The blue curve is a fitted mixture model and the red curves are scaled, normal mixture components with means of 0.40 and 0.70, respectively.

PRELIMINARY DATA: DNA Methylation in Pacific oysters

Given the diversity of DNA methylation across taxa, the novel perspective offered by recent studies on invertebrates, and the limited knowledge of epigenetic mechanisms in molluscs, our lab set out to characterize DNA methylation in the Pacific oyster, *Crassostrea gigas*. ***This species is an excellent system to examine the interface of epigenetic mechanisms and environmental change because of the significant amount of genomic resources available and life history characteristics (i.e. intertidal, planktonic larvae, filter-feeding) that make it an effective bioindicator species***. In this section we will discuss the progression of our research efforts from initial evidence of DNA methylation in this species to current efforts sequencing methylated portions of the genome.

Initial bioinformatic analysis revealed the presence of genes that encode DNA methyltransferases (DNMT) and methyl-binding proteins (MBD) in the oyster genome.

Specifically, sequences with high homology to DNMT3, DNMT1, and MBD2 are present (Gavery and Roberts 2010). The use of methylation sensitive restriction enzymes and bisulfite sequencing provided direct experimental evidence of DNA methylation in the oyster (Gavery and Roberts 2010).

To begin to investigate the functional role of DNA methylation, we used the ratio of observed to expected CpG dinucleotides (CpG O/E) (described above) to predict methylation status in the *C. gigas* transcriptome. A majority of *C. gigas* genes analyzed were depleted in CpG dinucleotides (*i.e.* CpG O/E <1.0) indicating that DNA methylation is a common feature of the *C. gigas* transcriptome (**Figure 1**). Additionally, a significant difference in predicted methylation status was observed across functional gene families. Specifically, genes predicted to be hyper-methylated are ubiquitously expressed, critical genes such as those involved in DNA and RNA metabolism and ***genes predicted to be sparsely methylated (i.e. higher CpG O/E) are associated with tissue specific and inducible genes,*** including those involved in general immune function (cell adhesion, cell-cell signaling, and signal transduction). **Figure 2** shows the relationship between CpG O/E and biological process on the x-axis. A majority (65%) of the pair-wise comparisons of CpG O/E for each biological process are significantly different from each other (Gavery and Roberts 2010).

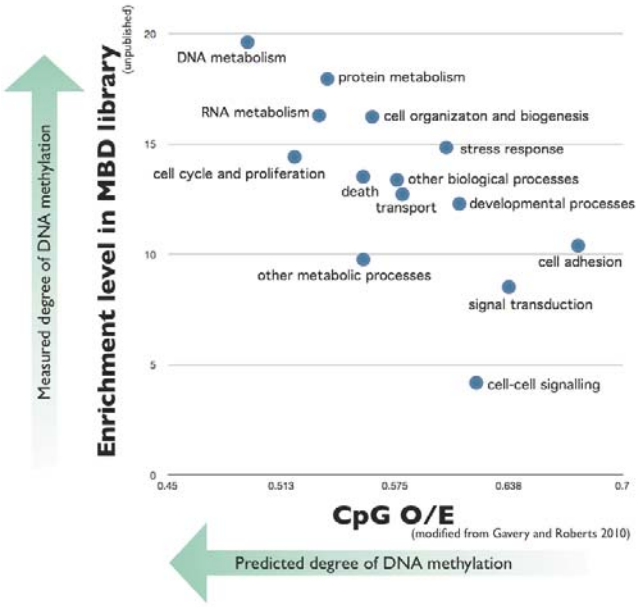

Figure 2. Predicted methylation level of *C. gigas* genes categorized by biological processes compared to experimentally measured level of DNA methylation. Mean CpG O/E for 10,699 *C. gigas* genes are plotted on the x-axis (modified from Gavery and Roberts 2010). DNA methylation was measured by constructing and sequencing a DNA library enriched for methylated DNA using the methyl-CpG binding domain of human MBD protein. .

We have experimentally characterized DNA methylation at the nucleotide level. Bisulfite sequencing was used to examine DNA methylation in *C. gigas* heat shock cognate 70 (Cghsc70) (GenBank Accession number AJ305315). We have sequenced approximately 1000 bp of the gene in an individual oyster gill sample. Methylation is primarily found in exons in the oyster, similar to what has been found in insects (Lyko et al 2010, Park et al 2011). For the two exons sequenced, 44% and 85% of the CpG's respectively, were determined to be methylated (**Figure 3**). For each CpG loci there is variability in the methylation status with anywhere between 0 - 100% of the clones showing methylation at each site.

We recently completed sequencing of methylation enriched DNA libraries using the SOLiD 4 high-throughput sequencing platform (Applied Biosystems). To generate the methylation enriched library, genomic DNA was randomly sheared and methylated fragments enriched by preferential binding to the methyl-CpG binding domain of human MBD2 protein prior to library preparation. The sequencing reads were mapped back to the *C. gigas* transcriptome. Transcriptomic resources are maintained and available via the online resource,

GigasDatabase (Fleury et al. 2009). Genes identified in the library enriched for methylated DNA were associated with their respective Gene Ontology (GO) terms and analyzed for enrichment based on the entire transcriptome (**Figure 2**). Results indicate the most underrepresented group in the methylated fraction is cell adhesion and the most enriched genes in this fraction are involved in DNA metabolism. These preliminary analysis are consistent with the results of the *in silico* analysis.

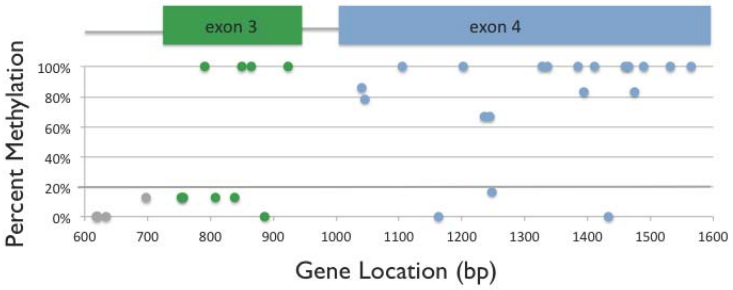

Figure 3. Methylation status of Cghsc70 as determined by bisulfite sequencing. Each circle represents a CpG loci. Percent methylation refers to number of clones analyzed. In all cases at least 8 clones were examined.

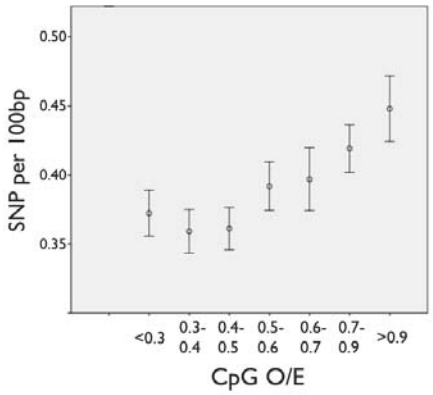

Figure 4. Sequence conservation as determined by mean SNPs per 100bp for genes grouped based on CpG O/E.

In an effort to determine the relative sequence conservation in genes methylated at the germline compared to those that are not, we mapped an RNA-seq library from pooled gill tissue and determined the number of single nucleotide polymorphisms (SNPs) (**Figure 4**). This analysis shows that genes predicted to be methylated at the germline (*i.e.* low CpG O/E) have less genetic diversity compared to genes lacking germline methylation (*i.e.* high CpG O/E). This is consistent to results from Lyko et al. (2010) where they showed increased sequence conservation in low CpG O/E genes across species. *To our knowledge, while preliminary, this is the first time a relationship between genetic diversity and CpG O/E among individuals in a population has been described.*

DNA methylation as a mechanism to increase adaptive potential

DNA methylation is associated with transcriptional activity and it likely functions in a complex nature to influence phenotypic diversity. In this research proposal we set out to test our working theory that in species subjected to highly heterogeneous environmental conditions, the DNA methylation system assists in providing increased adaptive potential. This theory is based on 1) the species' life history, 2) prevalence of DNA methylation in exons, and 3) the distribution of DNA methylation based on gene function.

The working theory, which will be referred to as the ***Methylation Enhanced Random Variation Theory, suggests that the absence of germline methylation in genes involved in effectively adapting to heterogeneous conditions, facilitates random variation that contributes to phenotypic plasticity and increased adaptive potential.*** In other words, in species that can be exposed to a wide variety of selective pressure, the DNA methylation system has evolved to increase survival. It is expected that genes that lack germline methylation (*i.e.* high CpG O/E genes) will have a larger number of "transcriptional opportunities" as compared to the ubiquitously expressed, critical genes. One "opportunity"

or variation that is expected is the production of alternative transcripts that might arise from 1) alternative start sites, 2) alternative stop codons, and/or 3) alternative combinations of exons. We also expect that there will be more sequence variation in this suite of genes, which could contribute to increased phenotypic plasticity. In addition, there is the opportunity for these genes to be transiently methylated, which would also influence transcription. Conversely, germline methylation limits mutation and alternative splicing in critical genes that are core to survival. **Figure 5** provides a schematic representation of this working theory.

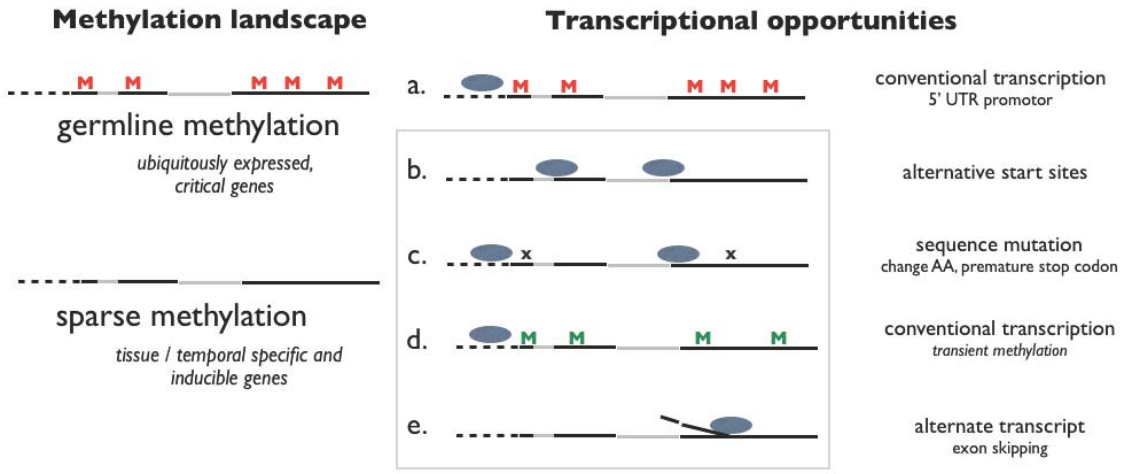

Figure 5. Schematic representation of the working theory on how DNA methylation influences transcriptional activity. The *Methylation Enhanced Random Variation Theory* proposes the absence of germline methylation contributes to adaptive potential. Transcriptional opportunities are diagramed for genes possessing germline methylation (a) and genes lacking methylation (b-e). Dashed lines represent the 5' UTR, solid lines represent exons and grey lines indicate introns. "M" designates a methylated CpG. "x" represents a sequence mutation. Blue ovals represent promoter complexes.

The *Methylation Enhanced Random Variation Theory* is consistent to what has been described in honey bees by Lyko et al. (2010) suggesting that methylation could "control which of several versions of a gene is expressed". Our working theory is also consistent with the relationship between methylation and phenotypic plasticity in plants. Early work examining DNA hypomethylation mutants in plants revealed dramatic differences in developmental phenotypes, indicating DNA methylation could suppress plasticity (Kakutani et al. 1996). Similarly, Bossdorf et al. (2010) found that treatment of *A. thaliana* with 5-azaC, a demethylation agent, increased phenotypic variability in several traits. The epigenetic variation found in *A. thaliana* was independent of genetic variation (Bossdorf et al. 2010). While other researchers have shown a relationship between DNA methylation, alternative splicing, and sequence conservation (*e.g.* Lyko et al. 2010, Park et al. 2011) and suggested a role for DNA methylation in influencing ecologically important traits (*e.g.* Angers et al. 2010), ***this working theory is the first inclusive framework that offers a suite of specific, testable mechanisms that contribute to evolutionary success by increasing the number of phenotypes via regional, random variation.***

The absence of DNA methylation in genes that are induced in response to changing conditions could allow for multiple transcripts indirectly by providing access to alternative promoter sites. This explanation is consistent with the ability of DNA methylation to inhibit binding of transcription factors to response elements in mammalian promoter regions (Iguchi-Ariga and Schaffner 1989). A recent mammalian study provided direct evidence of this, revealing that intragenic methylation limits the generation of alternate gene transcripts by masking intragenic promoters (Maunakea et al. 2010). Direct evidence of DNA methylation associated with alternative transcripts is also available in invertebrates. In the honey bee, the gene GB18602 has two forms (long and short) which were distinguished by an exon being skipped in the long form (Lyko et al. 2010). This exon contained a stop codon that creates a shorter, alternative transcript. The researchers went on to find numerous examples of genes where the methylated CpGs were associated with differentially spliced exons (Lyko et al. 2010). In our working theory this phenomenon would be consistent with the transient (or differential) methylation that could lead to alternative transcripts under different environmental conditions.

Sequence variation is another important source of potential phenotypic variation. There are several instances of evidence supporting a relationship between decreased methylation and increased sequence variation. A study in *Nasonia* showed high CpG O/E ratios correspond with higher substitution rates between related species for synonomous, nonsynomymous and intron sites (Park et al. 2011). In other words, there was more genetic variation in genes lacking germline methylation. This is further supported by Lyko et al. (2010) where evaluation of sequence similarity comparisons across species showed increased sequence conservation in methylated portions of the genome. Our lab has carried out similar analysis in *Crassostrea gigas* by comparing CpG O/E to SNPs/bp in the oyster transcriptome which indicates greater genetic variation in genes lacking germline methylation (**Figure 4**).

In summary, the Methylation Enhanced Random Variation Theory proposes that the absence of germline methylation affords the organism an increased adaptive potential by facilitating random variation in a portion of the genome responsible for maintaining homeostasis under selective pressure. The mechanisms described under this theory may only pertain to select lineages. These mechanisms would be advantageous in species such as marine invertebrates, where planktonic larvae are at the mercy of the currents and adults live in fluctuating, heterogeneous environments. On a larger time scale, DNA methylation will likely play a significant role in the ability of species to respond to global climate change by increasing the probability of successful adaptation compared to the expectations based on conventional genetic theory alone. ***Thus, a better understanding of the phenomenon and evaluation of this working theory will not only provide important information on molecular processes but will also improve our ability to predict ecosystem responses from an evolutionary perspective.***

Other characteristics of *C. gigas* that are consistent with this theory are the inherent sequence mutations and incidences of alternatively spliced transcripts. One dramatic example with respect to sequence variation in *C. gigas* was described by Taris et al. (2008), where they demonstrated that statistically significant different quantitative PCR data could be obtained using different primers, simply due to sequence variation that happened to correspond to the primer binding site.

There are numerous examples of alternatively spliced products in *C. gigas* and closely related species. Multiple subtypes of a gonadotropin-releasing hormone receptor exist in the Pacific oyster, each with different ligand specificity and function (Rodet et al. 2008). Recently, Kawabe and Yokoyama (2011) found that heat shock transcription factor 1 (HSF1), which regulates several genes involved in the stress response including heat shock proteins,

has eight alternatively spliced isoforms. Interestingly, there are multiple HSF encoded by several genes in vertebrates (Pirkkala et al., 2001). In oysters, only a single gene is found and thus the eight isoforms could make up the diversity that is lacking from having only one HSF gene.

The actual relationship between DNA methylation and phenotype is likely more complex than our working theory details. For instance, an alternative theory could involve a direct relationship where environmental conditions elicit a change in DNA methylation that corresponds to a specific, advantageous transcriptional response. Regardless, the *Methylation Enhanced Random Variation Theory* provides a framework for us to design a research effort that will provide significant insight into how DNA methylation influences transcriptional activity.

PLAN OF WORK

The aim of the current proposal is to elucidate the mechanisms by which DNA methylation regulates transcriptional activity. Based on other research and our own lab's work, we know that DNA methylation is an ancient process that has evolved to carry out different roles in different lineages. In a marjority of invertebrate species where DNA methylation has been examined, distinct characteristics emerge including, a) germline DNA methylation is primarily located in exonic regions, b) germline DNA methylation is found predominantly in ubiquitously expressed, critical genes with limited methylation in inducible genes, and c) DNA methylation is a dynamic process. Based on this information we have developed a working theory on the role and mechanisms by which DNA methylation relates to transcriptional activity. The specific research objectives of the current proposal are to 1) experimentally characterize the DNA methylation landscape in *Crassostrea gigas* and 2) evaluate the relationship between methylation, expression, alternative splicing, and sequence mutations.

Research Objective 1: Characterize the DNA methylation landscape in *Crassostrea gigas*

The first research objective expands on our *in silico* analysis and preliminary MBD-seq analysis. Based on what has been observed in other invertebrates and the relationship of CpG O/E across species we expect the same relationship between methylation and functional classification of genes. In addition, we will explore the degree of variation in DNA methylation patterns for this species. The specific hypotheses associated with this research objective include:

Hypothesis 1a: *DNA methylation will be predominantly found in ubiquitously expressed genes.*

Hypothesis 1b: *DNA methylation patterns will not significantly vary during development.*

In order to test these hypotheses, the methylome of sperm and larvae will be determined using MBD-seq. Adult oysters will be collected from the School of Aquatic and Fishery Sciences field site located in Puget Sound, Washington. Oysters will be maintained in the School of Aquatic and Fishery Sciences Aquatic Organism Facility under ambient conditions. Following two weeks of conditioning with an enriched microalgae diet, oysters will be strip spawned where sperm from at least 4 males will be added separately to sterile containers to fertilize eggs from an individual female. At the time of spawning, sperm will be immediately frozen at -80ºC for later DNA analysis. Oyster larvae will be raised under ambient conditions and sampled at 5 days post fertilization for later DNA analysis. The lab of the PI has experience in spawning shellfish and raising larvae. DNA will be extracted from sperm samples of individual males and from pools of larvae that correspond to paternal samples. The sperm sample was selected for analysis analysis because it offers a single cell type, which will

allow us to remove any potential bias of cell type specific methylation patterns. The larvae sample was selected for five days post fertilization because there are significant changes in tissue specific gene expression occurring. The vertical integration of sampling also provides the potential to qualitatively assess if DNA methylation patterns (assuming they vary across individuals) are heritable. Sampling multiple individuals will provide insight into the degree of natural epigenetic variation.

Methylation enrichment will be performed for each sample using the Methyl-Miner Kit (Invitrogen). This system enriches fragmented genomic DNA samples by preferential binding of methylated fragments to the methyl-CpG binding domain of human MBD2. This is the same approach that was used for our preliminary MBD-Seq analysis (**Figure 2**). Library construction and sequencing will be carried out using the Illumina HiSeq 2000 platform at the Fred Hutchinson Cancer Research Center Shared Resources core facility (see Letter of Collaboration). Sequencing reads that represent the methylated portions of the genome will be assembled on a genomic scaffold to determine how methylated portions of the genome correspond to annotated genes (i.e. coding sequences) and to make comparisons across samples. At the time of proposal submission, a completed fully assembled genome is not available, however, there are significant genomic resources for this species that will aid in analysis. Currently, the Pacific oyster transcriptome is assembled, annotated, and accessible online (GigasDatabase). Additional sequence information will come from other research efforts in the lab of the PI (SRA Accession #SRA043778), BAC clones, and genomic short read sequencing efforts. The latter includes 97.6 gigabases of genomic sequence recently released from the genomics institute BGI in Shenzhen, China (SRA043580). This is associated with the draft release of the oyster genome, as BGI stated in a press release that they had sequenced the Pacific oyster genome (Nature News, 2010). Analysis will be carried out by mapping quality trimmed reads back to the common genomic scaffold so comparisons can be made with respect to distribution of methylation marks and to determine if specific functional classes of genes are methylated. Analysis will be conducted using CLC Genomics Server running on hardware in the lab of the PI as well as Galaxy (Goecks et al. 2010; Blankenberg et al. 2010). If interesting trends are identified when comparing generations, we will consider performing limited MBD-chip analysis (see following section for a description).

Research Objective 2: Evaluate the relationship between methylation, expression, alternative splicing, and sequence mutations

Completion of the second research objective will provide important insight into the actual mechanism(s) by which DNA methylation impacts transcriptional processes. Based on the Methylation Enhanced Random Variation Theory, the absence of DNA methylation is a mechanism contributing to phenotypic plasticity. The hypotheses to be tested include:

Hypothesis 2a: *Differential methylation patterns will occur in conjunction with a change in environmental conditions.*

Hypothesis 2b: *Genes not methylated at the germline will have more alternative transcriptional products compared to genes with germline methylation.*

The first hypothesis is based on the idea that transient methylation will occur in genes lacking germline methylation to provide an additional transcriptional opportunity. In order to test both hypotheses, adult oysters (n=6) will be subjected to a temperature shock with gill tissue taken prior to and following the stressor for each individual in order to characterize methylation and transcriptional properties. RNA-seq will be performed on each sample independently using the Illumina HiSeq platform and barcoding of libraries. RNA-seq will allow us to evaluate expression, alternatively spliced products and sequence variation. ***DNA tiling arrays will be used to assess differential methylation in oysters pre- and post-stress***.

Specifically, a 3 x 720k chip will be designed by NimbleGen Roche with probe sizes ranging from 50 – 75 bp in length. The array will be designed in the latter part of year 1 and will use genome sequence to design targeted probes approximately every 100 bp across the genome. ***Even if this effort were to be carried out today, the available transcriptome and draft whole genome data would allow for the production of a chip that would be an exceptional tool to examine methylation***. The resolution of the chip will allow us to determine if individual exons are differentially methylated, which will provide clues to the functional significance of specific methylation marks (*i.e.* contributing to alternatively spliced products).

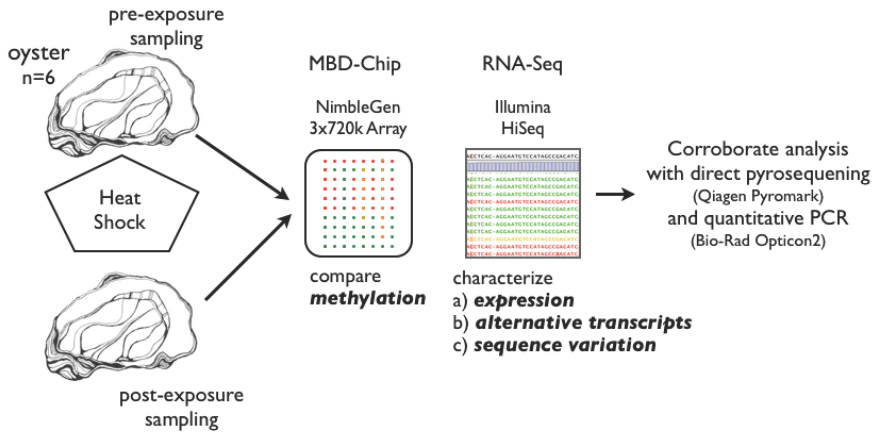

Figure 6. Schematic outlining experimental procedures associated with the completion of Research Objective 2. Oysters will be subjected to a temperature shock that will induce changes in gene expression. Individual oysters will be examined to analyze the relationship between methylation, expression, alternative splicing, and sequence mutations.

The MBD enrichment is the same procedure used for MBD-seq library construction component of Research Objective 1. The tiling array based approach, referred to as MBD-chip, was selected (as opposed to a direct sequencing based method) because the array platform offers an efficient and economical means to compare genome-wide methylation changes associated with different treatments or conditions. For instance, this technology has been used to analyze methylation patterns in cancerous cells compared to controls (Yegnasubramanian et al. 2011). In addition, methylation enriched tiling arrays have been used to evaluate whole organism exposures to environmental compounds such as cocaine, BPA, and pesticides (Novikova et al. 2008, Yaoi et al. 2008, Guerrero-Bosagna et al. 2010).

Oysters will be collected and maintained as described in Research Objective 1. The experimental design is intended to be able to directly compare methylation and transcriptional activity in a single individual to minimize confounding bias of any inherent genotypic or epigenetic variation. The type of stressor is not critical, only that it induces changes in gene expression. A temperature challenge was selected because it has been shown previously to elicit a transcriptomic response (Meistertzheim et al. 2007, Farcy et al. 2009). Furthermore, the anatomy of an oyster allows for non-lethal sampling of the responsive cells (*i.e.* gill tissue). It is expected that the sampling procedure itself will also impact the transcriptional activity.

Oysters (n=6) will be initially sampled for both RNA and DNA by notching the shell edge and removing gill tissue. The gill tissue was selected because this tissue is rich in hemocytes, the primary immune cells. Following tissue sampling, water temperature will be raised 10 degrees Celsius over three hours. Research in the lab of the PI has previously demonstrated that this treatment will have a significant physiological effect. Gill tissue from each oyster will be sampled again post stress. DNA and RNA will be isolated from each

sample (n=12).

Pre- and post-stress samples will be co-hybridized on one array. Using the 3 x 720k slide design, methylation patterns of three oysters can be directly compared on a single slide. Analysis will be performed to determine if differential methylation patterns exist following the change in environmental conditions. Specifically, we expect methylation to occur primarily in exons and expect that methylation patterns will not change significantly in genes predicted to be methylated at the germline (ubiquitously expressed genes). For those genes that are differentially expressed and/or predicted to be associated with a physiological response (based on annotations), it is expected that there will not be a direct relationship between methylation and expression. In other words, based on the working theory that DNA methylation is a fundamental component of the genome that improves adaptive potential, there will not be a direct relationship between level of expression and methylation. This prediction is also supported by research in insects showing there was not a consistent change in local DNA methylation that accompanied differential gene expression (Lyko et al. 2010). Based on the Methylation Enhanced Random Variation Theory (**Figure 5-d**) it is expected that some transcripts will be transiently methylated. Details of DNA array analysis are provided in the Methods section below and will be carried out in conjunction with Fred Hutchinson Cancer Research Center Shared Resources core facility staff (see Letter of Collaboration).

While the MBD-chip approach is ideal for comparing samples it does not provide single base pair resolution of CpG methylation. Therefore, once differentially methylated regions have been identified, methylation will be interrogated directly by pyrosequencing bisulfite treated DNA from individual oysters. This is a quantitative sequencing-by-synthesis method that monitors real time incorporation of nucleotides through the enzymatic conversion of released pyrophosphate into a proportional light signal (Tost and Gut 2007). This will allow us to confirm differential methylation between samples and provide methylation status of individual CpG sites providing insight into the location and number of differentially methylated CpG sites required to induce a functional change.

RNA analysis will be carried out by the construction and sequencing of 12 cDNA libraries on the Illumina HiSeq 2000 platform. Total RNA will be extracted from gill tissue samples followed by two rounds of mRNA enrichment (MicroPoly(A) Purist, Ambion). Libraries will be constructed by Fred Hutchinson Cancer Research Center Shared Resources core facility staff (see Letter of Collaboration) and barcoded so multiple libraries can be multiplexed. Sequence reads will be quality trimmed and sorted. Sequence analysis will be performed using CLC Genomics Server (CLC Bio) and Galaxy. ***RNA-seq will allow us to evaluate gene expression,*** which compares the number of reads per gene across samples. The most recent assembly of the transcriptome (*GigasDatabase*) and genome will be used as the scaffold for mapping sequence reads. As indicated above, it is likely that there will be a significant amount of new genomic sequence information available that will be incorporated. The PI has significant experience using RNA-seq analysis to examine gene expression in non-model species including the oyster, abalone, coral, lake trout (Goetz et al. 2010), and herring (Roberts et al. *in review*).

A second transcriptome parameter that will be examined using RNA-seq is ***evidence for alternative transcripts.*** Based on the working theory (**Figure 5-b, 5-c, 5-e**) it is expected that genes lacking germline methylation will have a greater likelihood of multiple isoforms as a result of alternative start sites, premature stop codons, and exon skipping. These phenomena would facilitate localized random variation, increasing adaptive potential. For this analysis, the relative distribution of reads across a gene will be compared. For example, if for a given gene in one sample there was a predominance of reads mapping in the 3' region, it could

suggest an alternative downstream promoter. RNA-seq has been used effectively to identify both known and novel alternative splicing events (Mortazavi et al. 2008, Wang et al. 2008, Richard et al. 2010). In fact, RNA-seq provides more accurate identification of alternative isoforms than other methods such as exon arrays (Hiller et al. 2009). The lab of the PI has also used RNA-seq to begin to characterize alternative transcripts in the Pacific oyster and lake trout (unpublished data). These analyses will focus on genes that are expressed at high levels to increase resolution. Comparisons will be made, 1) between genes methylated at the germline and those that are not and 2) across individuals for the same gene. The specific methylation pattern at time of sampling (MBD-chip) will also be taken into consideration.

RNA-seq libraries will also be used to ***examine sequence variation***. Based on other research and our analysis of RNA-seq data from pooled individuals (**Figure 4**), it is expected that genes that are not methylated at the germline will have greater sequence variation. This analysis will be carried out by mapping reads to the genomic scaffold and performing single nucleotide polymorphism (SNP) and deletion/insertion polymorphism (DIP) analysis using CLC Genomics Server. The PI has successfully used RNA-seq data to identify sequence variation including SNPs in chum salmon (Seeb et al. 2010), herring (Roberts et al *in review),* and hard clam (unpublished). While it is expected that there will be a measurable difference in the number of SNPs at the species level, if in fact DNA methylation is directly involved in limiting mutations, we would expect to see genetic variation at the individual level. This level of analysis has not been previously carried out as the preliminary library was constructed from a pool of individuals. Similar to the analysis that will be performed to characterize alternative transcripts, comparisons will be made at the gene, functional gene category (*e.g.* GO annotations, CpG O/E), species, and individual level.

A limited number of targets will be selected and quantitative PCR performed as described in Roberts et al. (2008, 2011) to begin to examine the relationship between DNA methylation and transcriptional regulation. This will be used primarily to corroborate expression analysis and the presence of alternative transcripts. These combined data will allow us to directly compare DNA methylation and gene expression and provide important insight into the relationship with transcriptional activity and variation.

Specific Methods

Methylation Enrichment: Methylation enrichment will be performed using the Methyl-Miner Kit (Invitrogen) which binds fragmented double-stranded genomic DNA using biotin labeled Methyl Binding Domain 2 protein (MBD2), and can be eluted using a salt gradient. DNA will first be fragmented to ~150 bp, then incubated with MBD2 protein coupled to paramagnetic bead via a biotin linker. The bound fraction (methylated fraction) will be eluted with a high salt concentration buffer.

Array Design and Analysis: A custom tiling array will be designed utilizing NimbleGen Roche technology. A 3 x 720k design will be used with probes derived to cover the oyster genome. For each comparison, labeled samples will be co-hybridized to the array, washed, and scanned. Array images will be used for data extraction as pair files. Genomic feature format (GFF) files will be produced for visualization of scaled log2-ratio data. P-value files (.gff) files will be generated from the scaled log2-ratio data, where each probe is tested for positive enrichment of DNA methylation against all probes on the array. All array processing and analysis will be conducted at the Fred Hutchinson Cancer Research Center Core Facility at the University of Washington. This core facility has significant expertise with this platform.

Sequence Analysis: Sequencing data will be analyzed using CLC Genomics Server (CLC Bio) along with publicly available databases (NCBI, SWISS-PROT, GigasDatabase) and our own unpublished *C. gigas* RNA-Seq libraries. Analysis will include quality trimming, de novo

assembly, and BLAST (Altschul et al. 1990). Comparisons among libraries will be made within CLC Genomics Server (RNA-Seq) and Galaxy tools (http://main.g2.bx.psu.edu/). To identify differentially expressed genes, enriched biological themes and GO terms will be identified using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (Huang et al. 2009a, Huang et al. 2009b).

Quantitative PCR: Total RNA will be isolated from tissues using TriReagent (Molecular Research Center, Inc.) and DNase treated using the Turbo DNA-free Kit (Ambion) according to the manufacturer's "rigorous" protocol. Removal of genomic DNA will be confirmed via real-time PCR on DNAse-treated total RNA. RNA will be reverse transcribed using M-MLV Reverse Transcriptase (Promega) and oligo dT primers (Promega) according to the manufacturer's protocol. Real-time PCR will be performed using gene specific primers, designed using Geneious (Biomatters Ltd.), with 2x Sso Fast EvaGreenSuperMix (BioRad). Real-time PCR reactions will be run on an Opticon 2 Real-Time PCR System. Raw fluorescence data will be analyzed using Real-time PCR Miner (Zhao & Fernald 2005).

Pyrosequencing: DNA from individual oysters will be bisulfite treated using the Epitect Bisulfite conversion kit (Qiagen, Carlsbad, CA). Briefly, DNA will be subjected to treatment with sodium bisulfite at increased temperature to deaminate unmethylated cytosine residues to uracil. Following treatment, the solution will be desulfonated on a column, washed and eluted. Pyrosequencing will be carried out on the PyroMark Q24 located at the Fred Hutchinson Cancer Research Center core facility (see Letter of Collaboration).

PLAN OF WORK SUMMARY

Crassostrea gigas will be used as a model system to evaluate the working theory that the absence of germline methylation in genes involved in adapting to changing conditions has evolved to increase adaptive potential in heterogeneous environments. ***A combination of high-throughput sequencing and DNA tiling array analysis will be coupled with methylation enrichment to determine genome wide methylation patterns and assess how specific DNA methylation patterns influence transcriptional activity***. Multi-generational samples will be examined as well as individual oysters prior to and following an environmental perturbation. The experimental design will allow for a comprehensive characterization of mechanisms associated with DNA methylation including testing the theory that genes lacking germline methylation will possess a greater incidence of alternative transcripts, sequence variation, and transient methylation.

TRAINING, EDUCATION, AND OUTREACH

This research will directly involve the training of two graduate and multiple undergraduate students. Two graduate students will be participating on this project. Mackenzie Gavery is currently completing a masters degree in my lab and is starting a PhD program in Fall 2011. Mackenzie is the first author on the manuscript characterizing methylation in the oyster (Gavery and Roberts, 2010) and was recently awarded an EPA-STAR fellowship to study the impact of environmental contaminants on DNA methylation in oysters. Another student will be recruited to work on this project. The arrangement will allow Mackenzie, who has significant experience in these techniques, to train her peer as well as providing her opportunities to be trained in new approaches (MBD-chip). Undergraduates will also be directly involved in the project. I have mentored over 15 undergraduates that have carried out independent research projects in my lab in the past 4 years. ***I am dedicated to providing opportunities to minorities, and over 30% of these undergraduate students come***

from under-represented ethnic groups. Undergraduate students in the lab are able to shadow graduate students and staff and then work in collaboration with myself and a graduate student to develop and carryout independent research.

As is the case for other research in our lab, the results and opportunity to participate will be incorporated into the formal classroom environment. I teach several courses including Biology of Shellfishes (FISH310), Integrative Environmental Physiology (FISH441), and Bioinformatics for Environmental Sciences (FISH546). I routinely share ongoing research with all of my classes and the research proposed here would certainly be pertinent while exposing students in multiple stages of education (freshman - post graduates) to the research. Furthermore, the 400 and 500 level courses have lab components where the students participate in project development. In these cases, I am able to leverage our research lab activities to facilitate experimental procedures in the educational labs. For example, in FISH441 this past year, a group of students examined the effects of pathogens and heavy metals on oyster gene and protein expression, which complemented work we are involved with characterizing disease susceptibility in Pacific oysters. Currently we are in the final stages of preparing a manuscript detailing this project where these FISH441 students (undergraduates and graduates) will be primary authors.

Dissemination of our research activities outside of the University is also critical. Our lab is dedicated to outreach and practices open notebook science. ***All lab members, including myself, maintain publicly accessible electronic notebooks where research activities and results are documented in near real-time*** (genefish.wikispaces.com). We also use the wiki-based platform to share basic laboratory protocols and procedures. Based on tracking web traffic from May to August 2011 we reach a large audience with visitors from over 20 countries including 23 US states. The lab also uses other web-based platforms such as Flickr, Twitter, YouTube, Tumblr, and Facebook to disseminate information and interact with scientists and the public.

Our lab is involved in hands-on outreach activities where we are able to share research with people of all ages and backgrounds (fish.washington.edu/genefish/robertslab/Outreach). Examples of activities current lab members are involved in include, volunteering at the Seattle Aquarium, the Making Connections program through the UW Women's Center (pairs high school student interested in STEM fields with mentors) and, the Seattle Expanding our Horizons program. Mackenzie Gavery, the graduate student who will be working directly on this project has served as a mentor with Girls in Engineering, Math and Sciences (GEMS) for three years and has worked with GEAR UP, a program encouraging high school students form low income and underserved communities to attend college. I have worked with elementary age students teaching them about shellfish biology as well as judging Science Fairs and have participated in several Centers for Ocean Sciences Education Excellence (COSEE) events. Together as an entire lab we participate in the Paws-On Science Weekend at the Pacific Science Center in Seattle, WA. At this event as well as most other activities we focus our presentations on the Physiological Responses of Oysters in Puget Sound (PROPS: tinyurl.com/UWprops) where we are able to effectively scale our conversations with all age groups - from filter feeding to the modification of DNA by the addition of methyl groups. The social, economic, and environmental prominence of the oyster in this region renders our research of interest to the public.

REFERENCES CITED

Ahuja N, Issa JP. Aging, methylation and cancer. *Histol Histopathol* 2000, 15:835-842.

Altschul SF, Gish W, Miller W, Myers E.W, Lipman DJ. Basic local alignment search tool *J Mol Biol* 1990, 215:403-410.

Angers B, Castonguay E, Massicotte R. Environmentally induced phenotypes and DNA methylation: how to deal with unpredictable conditions until the next generation and after. *Mol Ecol* 2010, 19:1283-1295.

Bestor TH. DNA methylation: How a bacterial immune function has evolved into a regulator of gene expression and genome structure in higher eukaryotes. *Phil Trans Royal Soc Lond* 1990, 326:179–187.

Bird AP, Taggart MH, Smith BA. Methylated and unmethylated DNA compartments in the sea urchin genome. *Cell* 1979, 17:889–901.

Bird AP, Taggart MH. Variable patterns of total DNA and rDNA methylation in animals. *Nucleic Acid Res* 1980 8:1485-1497.

Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. Galaxy: a web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol* 2010, Chapter 19:Unit 19.10.1-21.

Bossdorf O, Arcuri D, Richards CL, Pigliucci M. Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in Arabidopsis thaliana. *Evol Ecol* 2010, 24:541–553.

Cheung H, Lee T, Rennert O, Chan W. DNA methylation of cancer genome. *Birth Defects Res* 2009, 87:335–350.

Chinnusamy V, Zhu J. Epigenetic regulation of stress responses in plants. *Current Opin Plant Biol* 2009, 12:133-139.

Choi CS, Sano H: Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. *Mol Genet Genome* 2007, 277:589-600.

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W. Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* 1978, 274:775–780.

Elango N, Hunt BG, Goodisman MAD, Yi S. DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc Natl Acad Sci USA* 2009, 106:11206-11211.

Farcy E, Voiseux C, Lebel, JM, Fiévet B. Transcriptional expression levels of cell stress marker genes in the Pacific oyster *Crassostrea gigas* exposed to acute thermal stress. *Cell Stress Chaperon* 2009, 14:371–380.

Fleury E, Huvet A, Lelong C, de Lorgeril J, Boulo V, Gueguen Y, Bachère E, Tanguy A, Moraga D, Fabioux C, Lindeque P, Shaw J, Reinhardt R, Prunet R, Davey G, Lapègue S, Sauvage C, Corporeau C, Moal J, Gavory F, Wincker P, Moreews F, Klopp C, Mathieu M, Boudry P, Favrel B. Generation and analysis of a 29,745 unique expressed sequence tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. *BMC Genomics* 2009, 10:341. *http://www.biomedcentral.com/1471-2164/10/341*

Foret S, Kucharski R, Pittelkow Y, Lockett GA, Maleszka R. Epigenetic regulation of the honey bee transcriptome: unraveling the nature of methylated genes. *BMC Genomics* 2009, 10:472. *http://www.biomedcentral.com/1471-2164/10/472*

Gavery M, Roberts SB. DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). *BMC Genomics* 2010, 11:483. *http://www.biomedcentral.com/1471-2164/11/483*

Geyer K, Rodriguez-Lopez CM, Chalmers IW, Munshi SE, Truscott M, Heald J, Wlkinson MJ, Hoffmann KF. Cytosine methylation regulates oviposition in the pathogenic blood fluke *Schistosoma mansoni*. *Nature Commun* 2011, 2:424

Goecks J, Nekrutenko A, Taylor J, The Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* 2010, 1:R86. *http://genomebiology.com/content/11/8/R86*

Goetz F, Rosauer D, Sitar S, Goetz G, Simchick C, Roberts S, Johnson R, Murphy C, Bronte C, MacKenzie S. A genetic basis for the phenotypic differentiation between siscowet and lean lake trout (*Salvelinus namaycush*). *Mol Ecol* 2010, 19:176–196.

Gowher H, Leismann O, Jeltsch A. DNA of Drosophila melanogaster contains 5-methylcytosine. *EMBO J* 2000, 19:6918–6923.

Guerrero-Bosagna C, Settles M, Lucker BJ, Skinner MK. Epigenetic transgenerational actions of vinclozolin on promoter regions of the sperm epigenome. *PLoS ONE* 2010, 5: e13100. *http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0013100*

Hiller D, Jiang H, Xu W, Wong W: Identifiability of isoform deconvolution from junction arrays and RNA-Seq. *Bioinformatics* 2009, 25:3056-9.

Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res 2009, 37:1-13.*

Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protoc 2009, 4:44-57.*

Iguchi-Ariga SM, Schaffner W. CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev* 1989, 3:612–619.

Jablonka E, Lamb MJ. Epigenetic Inheritance and Evolution: The Lamarckian Dimension. Oxford: Oxford University Press. 1995.

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ. Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proc Natl Acad Sci U S A* 1996, 93:12406–12411.

Kawabe S, Yokoyama Y. Novel isoforms of heat shock transcription factor 1 are induced by hypoxia in the Pacific oyster *Crassostrea gigas*. *J Exp Zool A Comp Exp Biol* 2011, 315A: 394-407.

Kucharski R, Maleszka J, Foret S, Maleszka R. Nutritional control of reproductive status in honey bees via DNA methylation. *Science* 2008, 319:1827-1830.

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery J, Lee L, Ye Z, Ngo Q, Edsall L, Antosiewicz-Bouret J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker J. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 2009, 462:315-322.

Lyko F, Foret S, Kucharski R, Wolf S, Falckenhayn C, et al. The honeybee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol* 2010, 8: e1000506. *http://www.plosbiology.org/article/info:doi/10.1371/journal.pbio.1000506*

Lyko F, Ramsahoye BH, Jaenisch R. DNA methylation in Drosophila melanogaster. *Nature* 2000, 408:538-540.

Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, Turecki G, Delaney A, Varhol R, Thiessen N, Shchors K, Heine VM, Rowitch DH, Xing X, Fiore C, Schillebeeckx M, Jones SJM, Haussler D, Marra MA, Hirst M, Wang T, Costello JF. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 2010, 466:253–257.

Meaney MJ, Szyf M. Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome. *Dialogues Clin Neurosci* 2005, 7: 103–23.

Meistertzheim AL, Tanguy A, Moraga D, Thebault MT. Identification of differentially expressed genes of the Pacific oyster *Crassostrea gigas* exposed to prolonged thermal stress. *FEBS J* 2007, 274:6392-6402.

Mortazavi, A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 2008, 5:585–587.

News briefing: 6 – 12 August 2010, The Week in Science: Oyster Genome. *Nature* 2010, 466:800-801. *http://www.nature.com/news/2010/100811/full/466800a.html*

Novikova SI, He F, Bai J, Cutrufello NJ, Lidow MS, Undieh AS. Maternal cocaine administration in mice alters DNA methylation and gene expression in hippocampal neurons of neonatal and prepubertal offspring. *PLoS ONE* 2008, 3: e1919. *http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.000199*

Palmer LE, Rabinowicz PD, O'Shaughnessy AL, Balija VS, Nascimento LU, Dike S, de la Bastide M, Martienssen RA, McCombie WR. Maize genome sequencing by methylation filtration. *Science* 2003, 302:2115–2117.

Park J, Peng Z, Zeng J, Elango N, Park T, Wheeler D, Werren JH, Yi SV. Comparative analyses of DNA methylation and sequence evolution using *Nasoni*a genomes. *Mol Biol Evol* 2011, doi: 10.1093/molbev/msr168. *http://mbe.oxfordjournals.org/content/early/2011/06/20/molbev.msr168.abstract.*

Pirkkala L, Nykanen P, Sistonen L. Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. *FASEB* J 2001, 15:1118–1131.

Richard H, Schulz MH, Sultan M, Nürnberger A, Schrinner S, Balzereit D, Dagand E, Rasche A, Lehrach H, Vingron M, Haas SA, Yaspo ML. Prediction of alternative isoforms from exon expression levels in RNA-Seq experiments. *Nucleic Acids Res* 2010, 8:e112. *http://nar.oxfordjournals.org/content/38/10/e112.long*

Riggs AD, Xiong Z, Wang L, LeBon JM. Methylation dynamics, epigenetic fidelity and X chromosome structure. *Novartis Found Symp* 1998, 214:214-232.

Roberts SB, Goetz G, White S, Goetz F. Analysis of genes isolated from plated hemocytes of the Pacific oyster *Crassostrea gigas*. *Mar Biotchnol* 2008, 11:24-44.

Roberts SB, Hauser L, Seeb L, Seeb J. (*in review*) Development of genomic resources for Pacific herring through targeted transcriptome pyrosequencing. *PLOS one.*

Roberts SB, Sunila I, Wilkfors GH. Immune response and mechanical stress susceptibility in diseased oysters, *Crassostrea virginica*. *J Comp Physiol B* 2011, doi: 10.1007/s00360-011-0605-z.

Rodet F, Lelong C, Dubos MP, Favrel P. Alternative splicing of a single precursor mRNA generates two subtypes of Gonadotropin-Releasing Hormone receptor orthologues and their variants in the bivalve mollusc *Crassostrea gigas*. *Gene* 2008, 414:1–9.

Schorderet DF, Gartler SM. Analysis of CpG suppression in methylated and nonmethylated species. *Proc Natl Acad Sci USA* 1992, 89:957–961.

Seeb, JE, CE Pascal, ED Grau, LW Seeb, WD Templin, SB Roberts, and T Harkins. Transcriptome sequencing and high-resolution melt analysis advance SNP discovery in duplicated salmonids. *Mol Ecol Resour* 2010, 11:335-348.

Simmen MW, Bird, AP. Sequence analysis of transposable elements in the sea squirt, *Ciona intestinalis*. *Mol Biol Evol* 2000, 17:1685–1693.

Simpson VJ, Johnson TE, Hammen RF. *Caenorhabditis elegans* DNA does not contain 5-methylcytosine at any time during development or aging. *Nuc. Acids Res 1986,* 14:6711–6719.

Steward N, Ito M, Yamaguchi Y, Koizumi N, Sano H. Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. *J Biol Chem* 2002, 277:37741-37746.

Suzuki MM, Kerr ARW, De Sousa D, Bird A. CpG methylation is targeted to transcription units in an invertebrate genome. *Genome Res* 2007, 17:625–631.

Taris N, Lang RP, Camara MD. Sequence polymorphism can produce serious artefacts in real-time PCR assays: hard lessons from Pacific oysters. *BMC Genomics* 2008, 9:234. *http://www.biomedcentral.com/1471-2164/9/234*

Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protocols* 2007, 2:2265–2275.

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature* 2008, 456:470–476.

Wang Y, Jorda M, Jones PL, Maleszka R, Ling X, Robertson HM, Mizzen CA, Peinado MA, Robinson GE. Functional CpG methylation system in a social insect. *Science* 2006, 314:645–647.

Yaoi T, Itoh, Nakamura K, Ogi H, Fujiwara Y, Fushiki S. Genome-wide analysis of epigenomic alterations in fetal mouse forebrain after exposure to low doses of bisphenol A. *Biochem. Biophys Res Commun* 2008, 376:563–567.

Yegnasubramanian S, Wu Z, Haffnew MC, Esopi D, Aryee MJ, Badrinath R, He TL, Morgan JD, Carvalho B, Zheng Q, De Marzo AM, Irizarry RA, Nelson WG. Chromosome-wide mapping of DNA methylation patterns in normal and malignant prostate cells reveals pervasive methylation of gene-associated and conserved intergenic sequences. *BMC Genomics* 2011, 12:313. *http://www.biomedcentral.com/1471-2164/12/313*

Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* 2006, 126:1189–1201.

Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* 2006 39: 61–69.

BIOGRAPHICAL SKETCH

STEVEN BEYER ROBERTS

Assistant Professor; University of Washington
School of Aquatic and Fishery Sciences, Box 355020, Seattle WA 98195
Phone: 206-866-5141 - Email: sr320@uw.edu - Web: fish.washington.edu/genefish

Professional Preparation

BSc, North Carolina State University, 1997, Natural Resources
PhD, University of Notre Dame, 2002, Biological Sciences

Appointments

2006-present	Assistant Professor - University of Washington
2006-present	Adjunct Assistant Scientist - Marine Biological Laboratory, MA
2003-2006	Assistant Research Scientist - Marine Biological Laboratory, MA
2002-2003	Post-doctoral scholar – USDA NRI Fellowship Marine Biological Laboratory, MA

Publications - closely related

Roberts SB, Sunila I, Wikfors G. (2011) Immune response and mechanical stress susceptibility in diseased oysters, *Crassostrea virginica*. Journal of Comparative Physiology B. DOI: 10.1007/s00360-011-0605-z

Seeb JE, Pascal CE, Grau ED, Seeb LW, Templin WD, Harkins T, Roberts SB. (2010) Transcriptome sequencing and high-resolution melt analysis advance single nucleotide polymorphism discovery in duplicated salmonids. Molecular Ecology Resources. 11:335-348.

Gavery M* and Roberts SB. (2010) DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). BMC Genomics. 11:483. *http://www.biomedcentral.com/1471-2164/11/483*

Roberts SB, Goetz G, White S, Goetz F. (2009) Analysis of genes isolated from plated hemocytes of the Pacific oyster, *Crassostrea gigas*. Marine Biotechnology. 11:24-44.

Roberts SB, Gueguen Y, de Lorgeril J, Goetz F. (2008) Rapid accumulation of an interleukin 17 homolog transcript in *Crassostrea gigas* hemocytes following bacterial exposure. Developmental and Comparative Immunology. 32:1099-1104.

Publications - other significant works

Seeb JE, Carvalho G, Hauser L, Naish K, Roberts SB, Seeb LW. (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Molecular Ecology Resources. 11:1–8.

Goetz F, Rosauer D, Sitar S, Goetz G, Simchick C, Roberts SB, Johnson R, Murphy C, Bronte C, Mackenzie S. (2010) A genetic basis for the phenotypic differentiation between siscowet and lean lake trout (*Salvelinus namaycush*). Molecular Ecology. 19:176–196.

Mäthger L, Roberts SB, Hanlon R. (2010) Evidence for distributed light sensing in the skin of cuttlefish, *Sepia officinalis*. Biology Letters. 6:600-603.

Publications - other significant works continued

Lyons MM*, Smolowitz R, Dungan C, Roberts SB. (2006) Development of a real-time quantitative PCR assay for the hard clam pathogen, Quahog Parasite Unknown (QPX). Diseases of Aquatic Organisms. 72(1):45-52.

Roberts SB, McCauley LAR, Devlin RH, Goetz FW. (2004) Transgenic salmon over-expressing growth hormone exhibit decreased myostatin transcript and protein expression. Journal of Experimental Biology. 207(21):3741-3748.

** indicates student author*

Synergistic Activities (select)

Board of Directors: Pan-American Marine Biotechnology Association (2009-present)

Panel Member: NSF-Evolutionary Processes (2011), USDA-SBIR (2007)

Organizer: USDA WERA099: Broodstock Management, Genetics and Breeding Programs for Molluscan Shellfish. Feb 28 - March 1, 2010

Faculty Mentor: Ocean and Coastal Interdisciplinary Science (OACIS) GK-20 Program (2011-present)

Advocate for open notebook science and data sharing. The public can follow our research activities on Tumblr, Twitter, Facebook, YouTube, and Flickr. All students and staff maintain open access electronic lab notebooks and share resources using a wiki-based platform (fish.washington.edu/genefish).

Collaborators

Bassem Allam, Stony Brook University; Wade Carden, Stony Brook University; Ralph Elston, Aquatechnics; Carolyn Friedman, University of Washington; Frederick Goetz, University of Wisconsin; Yannick Gueguen, Ifremer-CNRS-Universite´ de Montpellie; Roger Hanlon, Marine Biological Laboratory; Lorenz Hauser, University of Washington; Claire Horner-Devine, University of Washington; Yuk-ting Lau, Stony Brook University; Scott Lindell, Marine Biological Laboratory; Julien de Lorgeril, Ifremer-CNRS-Universite´ de Montpellie; Maille Lyons, Old Dominion University; Jim Seeb, University of Washington; Lisa Seeb, University of Washington; Roxanna Smolowitz, Marine Biological Laboratory; Inke Sunila, Connecticut Department of Aquaculture; Joseph Vallino, Marine Biological Laboratory; Glenn VanBlaricom, University of Washington; Evan Ward, University of Connecticut; Gary Wikfors, NOAA.

Graduate and Postdoctoral Sponsors

Graduate Advisor: Frederick Goetz, University of Wisconsin

Postdoctoral Advisor: Frederick Goetz, University of Wisconsin

Thesis Advisor and Post Graduate-Scholar Sponsor

Mackenzie Gavery, Caroline Storer, Dave Metzger, Emma Timmins-Schiffman, Lisa Crosson.

Total Number: 5 graduate students

SUMMARY PROPOSAL BUDGET

YEAR 1

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of Washington

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Steven Roberts

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Steven Roberts - Assistant Professor**	0.00	0.00	1.00	**8,040**	
2.					
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	1.00	**8,040**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	2.00	0.00	0.00	**6,152**	
3. (**1**) GRADUATE STUDENTS				**21,372**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**35,564**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**7,695**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**43,259**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**1,000**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**23,000**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**1,500**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**15,588**	
TOTAL OTHER DIRECT COSTS				**40,088**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**84,347**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **MTDC (Rate: 54.0000, Base: 68759)**					
TOTAL INDIRECT COSTS (F&A)				**37,130**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**121,477**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				**121,477**	

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Steven Roberts

ORG. REP. NAME*
Cristin Schmitt

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET

YEAR 2

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of Washington

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Steven Roberts

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Steven Roberts - Assistant Professor**	0.00	0.00	2.00	**16,884**	
2.					
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	2.00	**16,884**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	2.00	0.00	0.00	**6,460**	
3. (**1**) GRADUATE STUDENTS				**16,830**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**40,174**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**9,473**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**49,647**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**1,500**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**18,000**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**2,000**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**11,691**	
TOTAL OTHER DIRECT COSTS				**31,691**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**82,838**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **MTDC (Rate: 54.5000, Base: 71147)**					
TOTAL INDIRECT COSTS (F&A)				**38,775**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**121,613**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				**121,613**	

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Steven Roberts

ORG. REP. NAME*
Cristin Schmitt

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET

Cumulative

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of Washington

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Steven Roberts

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Steven Roberts - Assistant Professor**	0.00	0.00	3.00	**24,924**	
2.					
3.					
4.					
5.					
6. ( ) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	3.00	**24,924**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**2**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	4.00	0.00	0.00	**12,612**	
3. (**2**) GRADUATE STUDENTS				**38,202**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**75,738**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**17,168**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**92,906**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**2,500**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**41,000**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**3,500**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**27,279**	
TOTAL OTHER DIRECT COSTS				**71,779**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**167,185**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE)					
TOTAL INDIRECT COSTS (F&A)				**75,905**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**243,090**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				**243,090**	

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Steven Roberts

ORG. REP. NAME*
Cristin Schmitt

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

C ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

BUDGET JUSTIFICATION

Salaries and Wages

For the completion of the research objectives outlined in the proposal we are requesting, over the 2 years, salary for the PI (Roberts), a graduate student, and a technician. The PI (Roberts) requests a total of 3 months of summer salary over the entire project. Dr. Roberts will be responsible for overall project management and training of students. Given the amount of sample processing and routine molecular procedures, a total of 2 months per year of support is requested for a research technician. The technician will also assist the graduate students funded by this project. Two graduate students will be involved in this project. Mackenzie Gavery, is beginning a PhD program and participated in the initial DNA methylation characterization in oysters. She was recently awarded an EPA-STAR research fellowship to examine the effects of xenobiotics on oyster physiology (including methylation) and no funds are requested to cover her stipend. A second graduate student will be recruited to work on this project that will be the core of their graduate thesis. Funds are requested to cover 4 quarters of support in year 1 and 3 quarters of support in year 2 for this student. The student will TA one quarter in year 2. The student will be responsible for experimental procedures, processing, and analysis.

Fringe Benefits

These monies are budgeted at rates set by the University of Washington: 27.2% for faculty, 16.1% for graduate students, and 33.6% for technicians.

Travel

$2500 is requested to partially fund travel by the PI and graduate students to national and international meetings to present research results. This includes participation in the Plant and Animal Genome Conference. A limited proportion of these funds will be used for local travel to collect oysters and to attend local education and outreach activities.

Materials and Supplies

During year 1, the primary research activities will include characterizing DNA methylation patterns in oysters using MBD-Seq (Research Objective 1) and RNA-Seq analysis to characterize transcriptional variation (Research Objective 2). Both the spawning event and exposure of oysters to temperature stress will be carried out in year 1. $7,000 is requested to perform MDB-Seq using the Illumina Hi-Seq platform. This includes funds to construct at least eight libraries and perform sequencing. Sequencing will be carried out at the Fred Hutchinson Cancer Research Center at the University of Washington (see Letter of Collaboration). $1500 is requested to purchase materials and kits (*e.g.* Methyl-Miner, Invitrogen) necessary to perform methylation enrichment prior to library construction. RNA-seq analysis will also be performed on samples from the temperature stress experiment using the same platform and will cost $12,000. This will include construction and sequencing of 12 libraries (6 individuals, pre and post exposure). The remaining $2,500 will be used to purchase supplies for routine molecular work including supplies for DNA and RNA isolation, lab consumables (*e.g.* 96-well plates, gels, covers, tips, agar plates, media), and molecular reagents and plasticware needed for sequencing.

Supplies and materials used during year 2 will primarily be used for methylation enrichment and tiling array development and processing associated with the temperature stress experiment. $6,000 will be used for this and will include cost of arrays and materials for labeling and

processing arrays. Probe construction, hybridization, washing and scanning will be carried out at the Fred Hutchinson Cancer Research Center at the University of Washington. Targets that are identified as differentially methylated using this approach will be further interrogated by single template pyrosequencing. $4000 is requested to cover costs associated with sample preparation (*i.e.* bisulfite treatment) and pyrosequencing. Quantitative PCR will be used to characterize gene expression levels including the presence of alternatively spliced products, with $6000 requested to cover these expenditures. Supplies will include molecular lab reagents (e.g. TriReagent, primers, DNA extraction kits, Amplitaq DNA Polymerase Kit, Perkin & Elmer; Quantitative PCR Master Mix kits, DNase). The remaining $2000 will cover costs of common molecular reagents and supplies.

Tuition

The proposed project will be the focus of the thesis of a graduate student. He/she will commit two years to this project and therefore funds are requested ($27,279) to cover cost of tuition.

Publication Costs

A total of $3500 is requested to cover cost of publications used for outreach and education as well as costs associated with publishing research results in peer-reviewed journals. It is the intent to publish in open-access scientific journals, which typically have higher fees.

Indirect Costs

Indirect costs are calculated as 54% of modified direct costs for year 1 and 54.5% of modified direct costs for year 2.

Current and Pending Support

(See GPG Section II.C.2.h for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Steven Roberts	Other agencies (including NSF) to which this proposal has been/will be submitted.

Support: ☒ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support
Project/Proposal Title: Threats to bivalve aquaculture and fisheries: the influence of emerging diseases and environmental change

Source of Support: NOAA
Total Award Amount: $ 243,000 Total Award Period Covered: 09/01/09 - 03/30/12
Location of Project: University of Washington
Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 1.00

Support: ☒ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support
Project/Proposal Title: Effects of ocean acidification on declining Puget Sound molluscan calcifiers

Source of Support: Washington Sea Grant
Total Award Amount: $ 478,092 Total Award Period Covered: 03/01/10 - 02/28/13
Location of Project: University of Washington
Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 0.50

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support
Project/Proposal Title: Epigenomic variation in the Pacific oyster (focuses on hybrid vigor)

Source of Support: USDA
Total Award Amount: $ 255,063 Total Award Period Covered: 10/01/11 - 09/30/13
Location of Project: University of Washington
Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 1.00

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support
Project/Proposal Title: DNA methylation as a mechanism to increase adaptive potential in invertebrates

Source of Support: NSF
Total Award Amount: $ 243,090 Total Award Period Covered: 03/01/12 - 02/28/15
Location of Project: University of Washington
Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 1.50

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support
Project/Proposal Title:

Source of Support:
Total Award Amount: $ Total Award Period Covered:
Location of Project:
Person-Months Per Year Committed to the Project. Cal: Acad: Summ:

*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

FACILITIES, EQUIPMENT & OTHER RESOURCES

FACILITIES: Identify the facilities to be used at each performance site listed and, as appropriate, indicate their capacities, pertinent capabilities, relative proximity, and extent of availability to the project. Use "Other" to describe the facilities at any other performance sites listed and at sites for field studies. USE additional pages as necessary.

Laboratory: The PI has research laboratory (800 sq. ft) in the School of Aquatic and Fishery Sciences in the College of the Environment at the University of Washington. The lab is equipped with routine laboratory furniture, plumbing, chemical and biohazard hoods, and other utilities. Specific

Clinical:

Animal: All animal holding and experimental procedures will be carried out in the School of Aquatic and Fishery Sciences Aquatic Organism Facility. This facility has several recirculating seawater systems with the ability to control temperature and flow. All effluent from the facility is

Computer: The PI maintains in his lab an Apple Xserve Computer Cluster (80 CPU; 2.26 GHz Intel Xeon) for large scale bioinformatic analysis. This system runs Genomics Server (CLCBio), BFX Cell (CLCBio), several open source software packages, and custom scripts for analysis of high-throughput sequencing

Office:

Other: In addition to the computing resources described above, the eScience Institute operates Hyak, a shared, high-performance computer cluster dedicated to research computing at the University of Washington. Currently Hyak includes more than 3,800 CPU cores and nearly 13TB of RAM. The PI is a participant in this project and has access to the system

MAJOR EQUIPMENT: List the most important items available for this project and, as appropriate identifying the location and pertinent capabilities of each.

Illumina HiSeq 2000 (Fred Hutchinson Cancer Research Center)
Axon Instruments GenePix 4000B microarray scanner (Fred Hutchinson Cancer Research Center)
Maui Hybridization Stations (Fred Hutchinson Cancer Research Center)
PyroMark Q24 (Fred Hutchinson Cancer Research Center)
Bio-Rad Opticon2 quantitative PCR machine (Roberts Lab)

OTHER RESOURCES: Provide any information describing the other resources available for the project. Identify support services such as consultant, secretarial, machine shop, and electronics shop, and the extent to which they will be available for the project. Include an explanation of any consortium/contractual arrangements with other organizations.

The Genomics Shared Resource located in the Fred Hutchinson Cancer Research Center has state-of-the art instrumentation and experienced staff who provide microarray, NextGen sequencing, and RNAi screening services. Technologies include: Affymetrix, Agilent, Illumina, and NimbleGen (microarrays) and Illumina HiSeq 2000 (NextGen sequencing). The Resource also has sophisticated data management and analysis resources, as well as the expertise to assist in data processing and analysis.

LABORATORY FACILITIES (continued):

equipment in the lab includes Sorval Super T21 highspeed centrifuge; Lab-Line benchtop orbital bacterial shaking incubator; Fisher Biotech hybridization incubator; 2-Boekel incubator shakers; 3 thermocyclers (MJ-PTC-200, PTC-100, PTC-150); 2-Road Runner Owl electrophoretic mingel systems for SDS PAGE; Trans-Blot SD-Biorad transfer system; Opticon2 Continuous Fluorescence Detection System (BioRad) 2-EC 600 and 1-EC 452 power supplies; Fisher 550 ultrasonic tissue processor; Nanodrop Spectrophotometer, Various horizontal gel systems for RNA/DNA gels; TL-2000 translinker; Thelco high performance above ambient incubators for bacteria and hybridization; Balances - Cahn C-30 microbalance; top loading balances (O-Haus); semi- analytical balance Mettler AC-100); microcentrifuges; So-Low ultralow freezer (U85-22); 2 - 45.0 ft3 sliding glass door refrigerators.

ANIMAL FACILITIES (continued):

chlorinated (10ppm for 24 hr) prior to release. A TB UV airsterilizer controls aerosolization of any pathogens and iodinated foot baths control movement of pathogens into or out of the facility.

COMPUTER FACILITIES (continued):

data (e.g. Illumina). The Cluster also includes Bioteam's iNquiry software for standalone BLAST and 200 other in-house genomic applications. This system is maintained by the PI (Roberts).

The Roberts lab currently has two PowerMac G4s, three MacMinis, two Linux computers, and two Windows computer that controls the Nanodrop and Opticon2. Computers are equipped with bioinformatic software (i.e. Genomics Server -CLCBio, Geneious, MacVector) and statistical software (SPSS). Medium scale bioinformatic analysis is carried out on a Mac Pro (two 3.0 GHz dual-core Intel Xeon) using various packages (BLAST, PHRAP, PHRED).

The University of Washington maintains a wired and wireless network for all students, faculty and staff. The School of Aquatic and Fishery Sciences has a full-time computer and networking specialist on staff to support faculty and student projects.

OTHER FACILITIES (continued):

DATA MANAGEMENT PLAN

Data Description

As part of this research effort a significant amount to data will be generated including short read sequencing information (MBD-Seq and RNA-Seq) and hybridization array data (MBD-Chip). Raw data from both MBD-Seq and RNA-Seq will include over 100 million short reads (50-100bp) and each run will include two files. One file will contain base pair information (csfasta) and a second, the corresponding sequence quality information (qual). These files will be the basis of analysis that will primarily be carried out in CLC Genomics Server, Galaxy, and custom scripts. The second major source of data will be that obtained from the tiling array. The raw data format in this instance will be images of the array scans. Other raw data will include single sequence information from direct pyrosequencing and fluorescent data from quantitative PCR analysis.

Access and Sharing

Raw data from the Illumina Hi-Seq and the slide scanner (both located at the Fred Hutchinson Cancer Research Center Shared Resources core facility) will be transferred to the lab of the PI where it will be used for analysis and archived. Within one month of acquiring the raw data it will be uploaded into the appropriate repository at NCBI. Specifically, the short read sequence data would be deposited in SRA and microarray data in GEO. Data will not be released immediately but this will provide additional data redundancy. Data will be released once the results are published or no later than three months after the project end date. Raw data from procedures including qPCR will be available in real-time on the Roberts Lab wiki where all lab notebooks are open to the public (genefish.wikispaces.com). In addition, limited amount of analyzed data (*i.e.* assembled and annotated genomic data) will be made available to the public through our lab wiki and other public webpages (*e.g*. http://pinboard.in/u:sr320/t:annotable/). These data are will be in non-propriety formats such as tab-delimited txt files. Limited analyzed data and workflow will also be made available via Galaxy.

Metadata

All raw data will include necessary and detailed metadata including stringently following of MIAME standards (Brazma et al. 2001) for array data. Personnel at the core facility will work with PI to insure all information is accurate and accessible (see Letter of Collaboration). While there is not a single common standard for short read sequence data, essential information including a description of the sample, library, sequencing method will be included in the SRA repository. Data tags will allow the data to be easily retrievable at NCBI.

Archiving and Backup

As described, raw data from short read sequencing and tiling array analysis will be uploaded to NCBI for archiving as well as providing access. Raw data will also be stored in two physical locations in the School of Aquatic Fishery Sciences at the University of Washington. The entire lab wiki is archived weekly in PDF format.

FRED HUTCHINSON
CANCER RESEARCH CENTER
A LIFE OF SCIENCE

August 25, 2011

Steven Roberts, PhD
University of Washington
School of Aquatic and Fishery Sciences
Box 355020
Seattle, WA 98105

Dear Dr. Roberts,

I am pleased to offer this letter of support for your NSF research proposal entitled, “DNA methylation as a mechanism to increase adaptive potential in invertebrates”. As I have indicated to you during the development and planning of this research proposal, the Genomics Resource here at the Fred Hutchinson Cancer Research Center has significant experience and expertise in next-generation sequencing platforms (e.g. Illumina’s Hi-Seq 2000) and microarry processing. As part of your proposed research, we will provide services including Illumina Hi-Seq library construction and sequencing, NimbleGen MBD array processing, and direct pyrosequencing for sequence based detection of methylation (PyroMark Q24). Furthermore, we will work with you to insure proper data analysis as well as assisting in data management, including the release of data in public repositories (i.e. NCBI’s GEO and SRA).

We look forward to working with you and your graduate students on this exciting project.

Sincerely,

Jeffrey J. Delrow, PhD
Director and Staff Scientist
Genomics Shared Resource